(12) UK Patent Application (19) GB (11) 2 298 501 (13) A

(43) Date of A Publication 04.09.1996

(21) Application No 9417807.6

(22) Date of Filing 05.09.1994

(71) Applicant(s)
Queen Mary & Westfield College

(Incorporated in the United Kingdom)

Mile End Road, LONDON, E1 4NS, United Kingdom

(72) Inventor(s)
Melvyn Slater

(74) Agent and/or Address for Service
Batchellor, Kirk & Co
2 Pear Tree Court, Farringdon Road, LONDON, EC1R 0DS, United Kingdom

(51) INT CL^6
G06F 3/00 15/18

(52) UK CL (Edition O)
G4A AKS
U1S S2425

(56) Documents Cited
US 5008946 A US 4894777 A

(58) Field of Search
UK CL (Edition N) G4A AKS AUXX
INT CL^6 G06F 3/00 9/44 15/18
Online: WPI, INSPEC

(54) Movement detection

(57) A method of detecting a walking movement by the user of a virtual reality system, in order to modify the virtual environment surrounding the user to reflect his apparent movement. A sensor is mounted on the head of the user, and an adaptive learning system (in hardware or software) is trained to recognise the pattern of head movements that correspond to a mimicked walking motion, so that the user can navigate the environment by "walking on the spot".

GB 2 298 501 A

2298501

"Virtual Reality Systems"

Background of the Invention

This invention relates to methods and apparatus for providing interfaces between human users, and computer systems, particularly, although not exclusively, adapted for use in so called "Virtual Reality" systems. In particular, the present invention provides a system for adaptively recognising patterns or sequences of movements by the user, in three dimensional space, and for interpreting them so as to provide suitable inputs to the virtual reality software.

The invention also has more general application, in other systems in which it is necessary to monitor a persons movements, and detect patterns in movements, which can then be used to control other processes, for example in remote control applications.

Interaction between the human and computer system in Virtual Reality is usually effected by electro-magnetic tracking devices, such as the Polhemus system. The human user wears a head mounted display (HMD), with a sensor on the top, and a receiver is able to provide 3 dimensional tracking data of the user's head movements, which are transmitted to the computer system. This is then used by the computer system to continually update the view of the scene presented to the user through the HMD. Similarly, the user holds a pointing device, or wears a data glove, which is used to transmit information about the position and orientation of the user's hand.

A problem is to provide a means for the user moving through the environment. Electro-magnetic sensors operate within a small field, so that it is not possible for participants to wander around a large Virtual Environment by physically walking. There have been several methods

implemented and discussed in the literature.

A standard solution for navigation in VR is to make use of the hand-held pointing device:

VPL used the DataGlove [FOLE 87]: a hand gesture would initiate movement, and the direction of movement would be controlled by the pointing direction. Velocity was controlled as part of the gesture: for example the smaller the angle between thumb and first finger the greater the velocity.

DIVISION's Pro Vision system typically employs a 3D mouse (though it supports gloves as well). Here the direction of movement is determined by gaze, and movement is caused when the user presses a button on the mouse. There are two speeds of travel controlled by a combination of button presses.

In order to give the user the sense of actually walking, rather than the artificial metaphor involved in using the hand, two solutions have been proposed:

Iwata [IWAT 92] used a system based on roller skates - the user would stand in a confined area wearing roller skates, and walk using the skates, while staying on the spot. The sensors on the skates would be used to return such foot movements to the computer, which could update the views accordingly.

Brooks [BROO 92] used a treadmill to the same effect - the user would walk on a treadmill, and the walking information transmitted to the computer to update the view.

These two solutions, although giving the user a more naturalistic sense of moving through the environment, require costly and cumbersome additional hardware.

Accordingly, a first aspect of the present invention provides a method of detecting a predetermined pattern of movement, in a system which exhibits a complex

relationship between the movements of its constituent parts, such as a human being in the process of walking, the method comprising the steps of:

(a) Determining the pattern of movement of one part of the system, over a period of time and

(b) Repeatedly applying the detected pattern, to an adaptive learning system, so as to progressively adapt it to produce a reliable indication of the subsequent occurrence of the predetermined pattern.

Preferably, the movement is detected in three dimensions, relative to a fixed point, and a series of x, y and z co-ordinates are supplied to the adaptive learning system in turn, so as to "train" it to recognise patterns which fall within a desired envelope of relationships between the values.

In one particular application of the invention, the system is used to recognise a human walking motion, by detecting the position of a sensor on the user's head. In this way, if the user mimics a walking motion, by "walking on the spot", the system can be "trained" to recognise the pattern of head movements which correspond to such walking, and thus, in a virtual reality system, the display seen by the user can be changed to reflect the changes which would be seen, if the user were actually walking in the "virtual space".

One possible embodiment of the "adaptive learning" system comprises a "neural network" which may comprise dedicated hardware circuitry or may be emulated in software.

Of course, it is also possible to detect movement by using a plurality of sensors, for example, one on each leg of the user. In practice, however, this is undesirable for a number of reasons:

(1) sensor devices of the required capabilities are expensive;

(2) if the number of signals to be processed is increased, the overall speed of operation of the system is reduced; and

(3) users are required to attach more pieces of equipment to themselves, which is inconvenient and can also be uncomfortable.

Thus the preferred arrangement, in which only one sensor is used, is greatly preferable.

It will also be appreciated that other types of "adaptive learning" systems could be used. For example, it may be possible to utilise a "evolutionary" program design, in which a program capable of recognising the desired pattern, is built up from a series of self modifying sub-routines, successive generations of which are more specifically adapted to the problem in hand. Other possible methods of pattern recognition could also be used, such as statistical method including discriminant analysis, or cluster analysis.

One embodiment of the invention will now be described in more detail, by way of example, with reference to a "virtual reality" system in which it is required to recognise that a user is "walking on the spot", so as to control the change of a display seen by the user, in accordance with his apparent movements through the virtual space.

2. Pattern Recognition: Recognising Walking on the Spot Behaviour in Virtual Reality

The method requires the detection of specific behavioural activity of users - that is, whether they are walking on the spot or doing something else. As an example, we have used a feed-forward neural net to implement a pattern recogniser that detects whether participants are "walking on the spot" or doing something else. However, there are other possible methods of pattern recognition, and artificial intelligence techniques such as genetic algorithms that would do the job. The neural network that has been implemented as a demonstrable example involves weighted back-propagation, so that changes detected in the weights are given a weight coefficient, depending on how great the change is. This is a standard method for pattern recognition described in [HERT 91].

The HMD tracker delivers a stream of position values (x_i,y_i,z_i) from which we compute first differences $(\Delta x_i,\Delta y_i,\Delta z_i)$. We choose a fixed sample of data $i = 1,2,\ldots,n$, and the corresponding delta-coordinates are inputs to the bottom layer of the net, so that there are 3n units at the bottom layer. There are two intermediate layers of m_1 and m_2 hidden units ($m_1 \leq m_2$), and the top layer consists of a single unit, which outputs either 1 corresponding to "walking on the spot" or 0 for anything else. We obtain training data from a person, which is used to compute the weights for the net. The network is then executed on the VR ProVision200 machine that we are using for all of our experiments.

After experimenting with a number of nets, we have found that a value of $n = 20$, $m_1 = 5$ and $m_2 = 10$ gives good results. We have never obtained 100% accuracy from any network, and this would not be expected.

There are two possible kinds of error, equivalent to Type I and Type II errors of statistical testing, where the null hypothesis is taken as "the person is not walking on the spot". The net may predict that the person is walking when they are not (Type I error) or may predict that the person is not walking when they are (Type II error). The Type I error is the one that causes the most confusion to people, and is also the one that is most difficult to rectify - in the sense that once they have been involuntarily moved from where they want to be, it is almost impossible to "undo" this. Hence our efforts have concentrated on reducing this kind of error. We do not use the output of the net directly, but only change from not moving to moving if a sequence of p 1s is observed, and from moving to not moving if a sequence of q 0s is observed ($q < p$). In practise we have used $p = 4$ and $q = 2$. The best result we have obtained is a correct

prediction 97% of the time. The Type I error is typically around 4% and the Type II error around 5%. It is likely that with further investigation of the Neural Net training method, results will improve.

The Polhemus Isotrak tracking device we are using actually returns data to the application at a rate of 28-30 Hz. The overall error is largely caused by the actual output lagging behind the real output by typically 5 samples, at the end of each sequence of 1s or 0s.

3. Incorporation into the Virtual Reality System

The following steps are carried out in order to support a person using this "walking on the spot" method. First, the person spends some time (typically 15 minutes) in the VR, where they are asked to "walk on the spot" some of the time, and do a range of other activities the rest of the time (eg, bending down, moving around, looking around, and so on). During this period, the data from the HMD that they are wearing is recorded, and segments of data (that is sequences of coordinate values) are marked as corresponding to "walking on the spot" or "other" activity. This exercise is carried out on the Provision200 Virtual Reality system.

The data is then transported to a SUN workstation, and a neural net training program is executed on the data. This is as described in Section 2. The data (that is sequences of first differences, each sequence marked as either "walking on the spot" or "other") is presented to the neural net trainer over and over again until the net "learns" - that is, an equation is established that for any data sequence, can predict whether or not this data sequence corresponds to "walking on the spot" or "other". The proportion of correct predictions give an indication of the success of the net. This is a standard method for training neural nets.

Once this equation is established, it is incorporated into the dVS software on the ProVision system, at the point in the software where the next view that the user will see is to be computed. At the moment where the next view is to be computed, the past sequence of HMD movements (ie. first differences of coordinates as described above) is put through the equation, and a prediction made, also based on previous predictions, as to whether the user is walking on the spot or not. If it is decided that the user is walking on the spot, then the view presented to them is computed based on the decision that they have moved forward. The direction of the move forward is determined by the direction of the user's gaze, which is also provided by the HMD.

The above paragraphs describe how a network is trained to fit the personal "walking on the spot" style of a user. In addition, we have arbitrarily designated the walking on the spot style of one person as "standard", and have trained other people to emulate this style, so that they are successfully able to move through the virtual environment relying on the trained network of this person.

4. Possible Benefits

We have carried out scientifically controlled experiments with users, comparing this "walking on the spot" method with the method that involves navigating with a hand-held pointing device (a DIVISION 3D mouse). The results are preliminary, since insufficient people have been through the experimental procedure at the time of writing. We have found that 9 out of 12 people have been able to use the "walking on the spot" method for moving through the environment - that is, the networks were sufficiently trained to correctly recognise their walking on the spot and "other" behaviour sufficiently often to allow them to successfully move around. The experiments have pointed out some problems with our data gathering procedures that we have rectified, so that we expect the proportion of successes to improve over time.

We have found that users probably find it easier and more accurate to move through the virtual environment by using the mouse. Also it is less tiring. However, there is some evidence to suggest that this result may be a function of the successful

performance of the neural network. For those users with networks that performed very well tended to score the "walking on the spot" method as being a preferable method of moving around than the method using the pointing device.

We have found that the walking on the spot method probably greatly enhances the person's sense of presence in the environment - that is the sense of "being there" in the computer generated environment rather than in the real world where their physical bodies were located. This is crucial, since it is the sense of presence that Virtual Reality system uniquely offer, so that anything which enhances the sense of presence is beneficial as a whole. In other words moving through the virtual environment using the mouse is probably perceived as less realistic by users, compared to walking on the spot.

5. Conclusions

A walking method for navigating through a virtual environment has been described. An example implementation was presented based on a neural network recognising a user's walking on the spot behaviour pattern, using data gathered from the tracking system for the HMD. Provisional results indicate that such a net can be successfully trained, and that this metaphor may enhance the sense of presence, in comparison to the more usual method of navigation using a hand-held pointing device. Of course, this is based on a small amount of data at the time of writing, though experiments are continuing.

It is less clear whether people prefer this walking method to using the mouse, purely from the point of view of actually getting around the environment. As Brooks (op. cit.) noted in the case of the real treadmill: "The steerable treadmill provided quite a realistic walking experience, and it neatly solved the problem of the limited range of the head sensor on the head-mounted display. Nevertheless, it proved to be too slow a tool for exploring extensive models. The user wore out with the exercise and grew frustrated at the slow pace. The flying metaphors proved more useful for this kind of rapid survey."

The utility of any metaphor depends on the application context. Clearly, just as in real life, walking is not a good method for exploring large spaces. It was observed that some of the experimental subjects did become physically tired as a result of walking, and it cannot be recommended to be used for a long time. However, it is a cheap additional tool in the range of interface metaphors available in VR, and there are circumstances where the sense of presence would outweigh the costs of relative inefficiency and tiredness. For example, consider an application for training simulation of emergency service personnel in hazardous conditions such as a fire: the fact that users would become tired and frustrated as a result of the additional exercise involved in a whole body movement is realistic. In real life, they would not move around a hazardous environment by using a mouse.

[BROO 92]
Brooks, F.P. et. al. (1992) Final Technical Report: Walkthrough Project, Six Generations of Building Walkthroughs. Department of Computer Science, University of North Carolina, Chapel Hill, N.C. 27599-3175.

[FOLE 87]
Foley, J.D. (1987) Interfaces for Advanced Computing, Scientific American, October, 126-135.

[HERT 91]
Hertz, J., A. Krogh, R.G. Palmer (1991) Introduction to the Theory of Neural Computation, Addison-Wesley Publishing Company.

[IWAT 92]
Iwata, H. and K. Matsuda (1992) Haptic Walkthrough Simulator: Its Design and Application to Studies on Cognitive Map, The Second International Conference on Artificial Reality and Tele-existence, ICAT 92, 185-192.

Claims:

1. A method of detecting a predetermined pattern of movement, in a system which exhibits a complex relationship between the movements of its constituent parts, such as a human being in the process of walking, the method comprising the steps of:
(a) determining the pattern of movement of one part of the system, over a period of time and
(b) repeatedly applying the detected pattern, to an adaptive learning system, so as to progressively adapt it to produce a reliable indication of the subsequent occurrence of the predetermined pattern.

2. A method according to claim 1 in which the pattern of movement is detected in three dimensions, relative to a fixed point, and a series of x, y and z co-ordinates are supplied in turn to the adaptive learning system, whereby it can be trained to recognise patterns which fall within a desired envelope of relationships between the values.

3. A method according to claim 1 or claim 2 in which the predetermined pattern of movement is a human walking movement, and the part being detected is the human head, whereby the occurrence of a walking motion can be detected from the corresponding movement of the head.

4. A virtual reality system in which a walking movement or mimicked walking movement by the user is detected by a method according to any of claims 1 to 3, whereby a display of a virtual environment surrounding the user can be modified in accordance with the user's movement or apparent movement.

5. A system according to claim 4 in which the said movement is detected by means of a sensor on the user's head.

6. A system according to claim 4 or claim 5 in which the adaptive learning system comprises a neural network which comprises dedicated hardware circuitry or a software emulation.

7. A method of detecting a predetermined pattern according to claim 1 and substantially as herein described.

8. A virtual reality system according to claim 4 and substantially as herein described.

The Patent Office

BEST AVAILABLE COPY

11

Application No: GB 9417807.6
Claims searched: 1 - 8

Examiner: Paul Nicholls
Date of search: 4 December 1995

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.N): G4A (AKS, AUXX)

Int Cl (Ed.6): G06F 3/00, 9/44, 15/18

Other: Online: WPI, INSPEC

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	US 5,008,946 A (ANDO) - Whole document	-
A	US 4,894,777 A (NEGISHI and HOSAAKA) - Whole document	-

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.